

14



OIIPE

RAW SEQUENCE LISTING

DATE: 05/05/2003

PATENT APPLICATION: US/09/927,811C

TIME: 14:22:47

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

3 <110> APPLICANT: RheinBiotech Limited Liability Company for New
 4 Biotechnological Processes and Products
 5 Romano, Ivano
 6 Gellissen, Gerd
 7 DeVergilio, Claudio

9 <120> TITLE OF INVENTION: HEAT-INDUCIBLE PROMOTER

11 <130> FILE REFERENCE: PCT1106-01966

13 <140> CURRENT APPLICATION NUMBER: US 09/927,811C

14 <141> CURRENT FILING DATE: 2001-08-09

16 <150> PRIOR APPLICATION NUMBER: PCT/EP00/01144

17 <151> PRIOR FILING DATE: 2000-02-11

19 <160> NUMBER OF SEQ ID NOS: 28

21 <170> SOFTWARE: PatentIn version 3.2

23 <210> SEQ ID NO: 1

24 <211> LENGTH: 792

25 <212> TYPE: DNA

26 <213> ORGANISM: Hansenula polymorpha

28 <400> SEQUENCE: 1

29 cttaaataacc acaataggaa aattatcaat aaagcttttc ggatttcatt acgttatatc 60
 31 gcaaaaaaat agtcgagctt tctgaaccgt tcgttaataa aaaaatagtt ttttcagatt 120
 33 tctatgtgag gcagtcacga tagaattcca tcgaactcgt cagcgccaaa tgtgaatgcg 180
 35 gctttcaaaa gctttgtcga atttgggatg ggaatccatg aatcgaagat gtcaaatg 240
 37 gggatcacaa aagtacactc acgaggaaaa tcaaacctt ctcgtacctt taacacatac 300
 39 ggaaatgatc gatcgatttg agaagattcc tcaatgattt tcgtcatata taggtatctg 360
 41 aggtatttat ggaccgattc gtaataacat catatacatc gcgctttgtc cctgtcccag 420
 43 agatttcgat gaaaaagcg aattttattc taatatttga agcatgccaa acatggggca 480
 45 gttgatttgt gtgagggtaa aatatcatga attgcaccca tcaaatgcag caagatattg 540
 47 accaatccta taatagaaaa cagacttacc acaaatagat tgtgatgacg atattatgaa 600
 49 tctccagatg aaaggctcga aagctatgaa gcctcttgaa acttttcatg gtgagataat 660
 51 attttcgaaa tttccacgaa cttctaaaac gcaattattg aatataaagg aaaaataata 720
 53 tttccatata gcaagcaaat caagctgcac tcctcatcct taaaactaat aaatcttacc 780
 55 catttgatac ca 792

58 <210> SEQ ID NO: 2

59 <211> LENGTH: 15

60 <212> TYPE: DNA

61 <213> ORGANISM: Artificial Sequence

63 <220> FEATURE:

64 <223> OTHER INFORMATION: Consensus sequence for a heat shock element

67 <220> FEATURE:

68 <221> NAME/KEY: misc_feature

69 <222> LOCATION: (1)..(15)

70 <223> OTHER INFORMATION: n may be a,c,t, or g

72 <400> SEQUENCE: 2

ENTERED

RAW SEQUENCE LISTING

DATE: 05/05/2003

PATENT APPLICATION: US/09/927,811C

TIME: 14:22:47

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

```

W--> 73 ngaannnnnnn ngaan 15
76 <210> SEQ ID NO: 3
77 <211> LENGTH: 15
78 <212> TYPE: DNA
79 <213> ORGANISM: Artificial Sequence
81 <220> FEATURE:
82 <223> OTHER INFORMATION: Special embodiment of the heat shock element
85 <220> FEATURE:
86 <221> NAME/KEY: misc_feature
87 <222> LOCATION: (1)..(15)
88 <223> OTHER INFORMATION: n is a, c, t, or g; b is g, c, or t; w is a or t; and m is c
or a
90 <400> SEQUENCE: 3
W--> 91 ngaannbwmm ngaan 15
94 <210> SEQ ID NO: 4
95 <211> LENGTH: 15
96 <212> TYPE: DNA
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: Nucleic acid sequence of a heat shock element
102 <400> SEQUENCE: 4
103 tgaagcctct tgaaa 15
106 <210> SEQ ID NO: 5
107 <211> LENGTH: 15
108 <212> TYPE: DNA
109 <213> ORGANISM: Artificial Sequence
111 <220> FEATURE:
112 <223> OTHER INFORMATION: Nucleic acid sequence of a heat shock element
114 <400> SEQUENCE: 5
115 tgaatataaa ggaaa 15
118 <210> SEQ ID NO: 6
119 <211> LENGTH: 1903
120 <212> TYPE: DNA
121 <213> ORGANISM: Hansenula polymorpha
123 <400> SEQUENCE: 6
124 atggtcaaag gtaatgttat agtggtttca aatagaatcc cagtcactat taagaagact 60
126 gaagatgatg aaaatggaaa atcaagatac gactatacaa tgtcatcagg cggattagtg 120
128 acggcattac aagggtctca aaatccattt cgatggtttg gatggcctgg gatgtctgtt 180
130 gatagcgaac agggacgaca aactgtcgag cgggatttga aggaaaagtt caattgttat 240
132 ccgatatggt taagtgcaga aattgcagac ttacattata acggcttttag caattctata 300
134 ctttgcccat tgttccacta tcacccaggg gagatgaatt ttgatgaaat tgcttgggcc 360
136 gcttattttg aagcaaataa actgttttgc caaacgatct taaaggagat aaaagacggg 420
138 gacgttatct ggggtacatga ttatcatctc atgttgttgc cttcactgct aagagaccaa 480
140 cttaatagta aggggtacc gaatgtcaaa attggctttt tccttcatac tccttttctt 540
142 tcaagcgaaa tatacaggat acttctgtga aggaaagaaa ttctcgaagg agtgcttagt 600
144 tgtgatttga taggtttcca cacctatgat tatgtccgtc actttcttag ttcggttgaa 660
146 agaattatga aattgcgaac gagcccacaa ggtgtgtctc ataatagatg acaggtgact 720
148 gtaagtgtct atccgattgg cattgacgtt gacaaattct tgaatggtct taagactgat 780
150 gaggtcaaaa gcaggataaa acagctggaa accagatttg gtaaagattg taaacttatt 840
152 attggggtgg acaggctgga ttacatcaaa ggtgtacctc aaaaactcca cgcgtttgaa 900

```

RAW SEQUENCE LISTING

DATE: 05/05/2003

PATENT APPLICATION: US/09/927,811C

TIME: 14:22:47

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

```

154 attttcttgg agagacaccc tgagtggatt ggaaaagttg ttttgatata ggtggctgtc 960
156 ccctcacgag gggacgttga agaatatcaa tctttgaggg cagctgtaaa tgagctagtg 1020
158 ggaagaatca atggttagatt tggtagcgtc gaatttggtc ctatccattt ccttcataaa 1080
160 agcgtgaact tccaagagct gatattctgtc tacgctgcta gtgatgtttg tgtagtgtca 1140
162 tcgacacggg acggaatgaa tttggtcagt tatgaatata ttgcttgtca acaagatcga 1200
164 aagggatctc tagtactaag tgaatttgcg ggagctgctc agtcattaata tggcgctctc 1260
166 gtagtgaatc catggaatac agaagaactc agtgaagcta tttacgaagg cttgatcatg 1320
168 agtgaagaga aaaggagggg caattttcag aagatgttca agtacattga gaaatatact 1380
170 gcaagttatt ggggagagaa ctttgtgaaa gaattgacga gagtgtgatt actgtggttt 1440
172 gcaggttaat ttgaaatgtt cacttgtact tgaagaattt tatattatat acatgttata 1500
174 catcaatagg ataaaaatta agtagacaaa gttatcattt tgttgggctg taaaaattga 1560
176 acgataacaa tatatttgac aaaaattaatt tgatctaatt gagctggagg gcgtaataata 1620
178 tttggtttcc tgaatcatct tgtagatcac aatatggggc agcttctttc gcagcgcgac 1680
180 acagagaaac acatcacact tgtccaacat gtcacatat cgcattcaat cggggaaatg 1740
182 caaggatata ggttgacat ggaagacgcy tctgtgatt tgaacgaaag aatattcgtg 1800
184 acggaagagg gacttgacat cagaaaacaa gacgagaata cagaggggtg tctggagtct 1860
186 cttcaaatta acatttatgg tgtctttgac ggacatggcg gtt 1903

```

189 <210> SEQ ID NO: 7

190 <211> LENGTH: 475

191 <212> TYPE: PRT

192 <213> ORGANISM: Hansenula polymorpha

194 <400> SEQUENCE: 7

```

196 Met Val Lys Gly Asn Val Ile Val Val Ser Asn Arg Ile Pro Val Thr
197 1 5 10 15
200 Ile Lys Lys Thr Glu Asp Asp Glu Asn Gly Lys Ser Arg Tyr Asp Tyr
201 20 25 30
204 Thr Met Ser Ser Gly Gly Leu Val Thr Ala Leu Gln Gly Leu Lys Asn
205 35 40 45
208 Pro Phe Arg Trp Phe Gly Trp Pro Gly Met Ser Val Asp Ser Glu Gln
209 50 55 60
212 Gly Arg Gln Thr Val Glu Arg Asp Leu Lys Glu Lys Phe Asn Cys Tyr
213 65 70 75 80
216 Pro Ile Trp Leu Ser Asp Glu Ile Ala Asp Leu His Tyr Asn Gly Phe
217 85 90 95
220 Ser Asn Ser Ile Leu Trp Pro Leu Phe His Tyr His Pro Gly Glu Met
221 100 105 110
224 Asn Phe Asp Glu Ile Ala Trp Ala Ala Tyr Leu Glu Ala Asn Lys Leu
225 115 120 125
228 Phe Cys Gln Thr Ile Leu Lys Glu Ile Lys Asp Gly Asp Val Ile Trp
229 130 135 140
232 Val His Asp Tyr His Leu Met Leu Leu Pro Ser Leu Leu Arg Asp Gln
233 145 150 155 160
236 Leu Asn Ser Lys Gly Leu Pro Asn Val Lys Ile Gly Phe Phe Leu His
237 165 170 175
240 Thr Pro Phe Pro Ser Ser Glu Ile Tyr Arg Ile Leu Pro Val Arg Lys
241 180 185 190
244 Glu Ile Leu Glu Gly Val Leu Ser Cys Asp Leu Ile Gly Phe His Thr
245 195 200 205
248 Tyr Asp Tyr Val Arg His Phe Leu Ser Ser Val Glu Arg Ile Leu Lys

```

RAW SEQUENCE LISTING

DATE: 05/05/2003

PATENT APPLICATION: US/09/927,811C

TIME: 14:22:47

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

```

249      210      215      220
252 Leu Arg Thr Ser Pro Gln Gly Val Val Tyr Asn Asp Arg Gln Val Thr
253 225      230      235      240
256 Val Ser Ala Tyr Pro Ile Gly Ile Asp Val Asp Lys Phe Leu Asn Gly
257      245      250      255
260 Leu Lys Thr Asp Glu Val Lys Ser Arg Ile Lys Gln Leu Glu Thr Arg
261      260      265      270
264 Phe Gly Lys Asp Cys Lys Leu Ile Ile Gly Val Asp Arg Leu Asp Tyr
265      275      280      285
268 Ile Lys Gly Val Pro Gln Lys Leu His Ala Phe Glu Ile Phe Leu Glu
269      290      295      300
272 Arg His Pro Glu Trp Ile Gly Lys Val Val Leu Ile Gln Val Ala Val
273 305      310      315      320
276 Pro Ser Arg Gly Asp Val Glu Glu Tyr Gln Ser Leu Arg Ala Ala Val
277      325      330      335
280 Asn Glu Leu Val Gly Arg Ile Asn Gly Arg Phe Gly Thr Val Glu Phe
281      340      345      350
284 Val Pro Ile His Phe Leu His Lys Ser Val Asn Phe Gln Glu Leu Ile
285      355      360      365
288 Ser Val Tyr Ala Ala Ser Asp Val Cys Val Val Ser Ser Thr Arg Asp
289      370      375      380
292 Gly Met Asn Leu Val Ser Tyr Glu Tyr Ile Ala Cys Gln Gln Asp Arg
293 385      390      395      400
296 Lys Gly Ser Leu Val Leu Ser Glu Phe Ala Gly Ala Ala Gln Ser Leu
297      405      410      415
300 Asn Gly Ala Leu Val Val Asn Pro Trp Asn Thr Glu Glu Leu Ser Glu
301      420      425      430
304 Ala Ile Tyr Glu Gly Leu Ile Met Ser Glu Glu Lys Arg Arg Gly Asn
305      435      440      445
308 Phe Gln Lys Met Phe Lys Tyr Ile Glu Lys Tyr Thr Ala Ser Tyr Trp
309      450      455      460
312 Gly Glu Asn Phe Val Lys Glu Leu Thr Arg Val
313 465      470      475
316 <210> SEQ ID NO: 8
317 <211> LENGTH: 2695
318 <212> TYPE: DNA
319 <213> ORGANISM: Hansenula polymorpha
321 <400> SEQUENCE: 8
322 cttaaataacc acaataggaa aattatcaat aaagcttttc ggatttcatt acgttatatc 60
324 gcaaaaaaat agtcgagctt tctgaaccgt tcgttaataa aaaaatagtt ttttcagatt 120
326 tctatgtgag gcagtcacga tagaattcca tcgaactcgt cagcgccaaa tgtgaatgcg 180
328 gctttcaaaa gctttgtcga atttgggatg ggaatccatg aatcgaagat gtcaaaatgg 240
330 gggatcacaa aagtacactc acgaggaaaa tcaaaacctt ctcgtacctt taacacatac 300
332 ggaaatgatc gatcgatttg agaagattcc tcaatgattt tcgtcatata taggtatctg 360
334 aggtatttat ggaccgattc gtaataacat catatacatc gcgctttgtc cctgtcccag 420
336 agatttcgat gaaaaaagcg aattttattc taatatttga agcatgccaa acatggggca 480
338 gttgatttgt gtgagggtaa aatatcatga attgcacca tcaaatgcag caagatattg 540
340 accaatccta taatagaaaa cagacttacc acaaatagat tgtgatgacg atattatgaa 600
342 tctccagatg aaaggctcga aagctatgaa gcctcttgaa acttttcatg gtgagataat 660

```

RAW SEQUENCE LISTING

DATE: 05/05/2003

PATENT APPLICATION: US/09/927,811C

TIME: 14:22:47

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

```

344 atttttcgaaa tttccacgaa cttctaaaac gcaattattg aatataaagg aaaaataata 720
346 tttccatata gcaagcaa at caagctgcac tcctcatcct taaaactaat aaatccttacc 780
348 catttgatac caatgggtcaa aggtaatgtt atagtgggtt caaatagaat cccagtcact 840
350 attaagaaga ctgaagatga tgaaaatgga aaatcaagat acgactatac aatgtcatca 900
352 ggcgggattag tgacggcatt acaagggtc aaaaatccat ttcgatgggtt tggatggcct 960
354 gggatgtctg ttgatagcga acaggagcga caaactgtcg agcgggattt gaaggaaaag 1020
356 ttcaattgtt atccgatatg gttaagtgc gaaattgcag acttacatta taacggcttt 1080
358 agcaattcta tactttggcc attgttccac tatcaccag gggagatgaa ttttgatgaa 1140
360 attgcttggg ccgcttattt ggaagcaa at aaactgtttt gccaaacgat cttaaaggag 1200
362 ataaaagacg gggacgttat ctgggtacat gattatcatc tcatgttgtt gccttcaactg 1260
364 ctaagagacc aacttaatag taagggtcta ccgaatgtca aaattggctt tttccttcat 1320
366 actccttttc cttcaagcga aatatacagg atacttcctg taaggaaaaga aattctcgaa 1380
368 ggagtgtcta gttgtgattt gataggtttc cacacctatg attatgtccg tcactttcctt 1440
370 agttcgggtt aaagaatatt gaaattgcga acgagcccac aagggtgttg ctataatgat 1500
372 agacagggtg ctgtaagtgc ttatccgatt ggcattgacg ttgacaaatt cttgaatggt 1560
374 cttaagactg atgaggtcaa aagcaggata aaacagctgg aaaccagatt tggtaaagat 1620
376 tgtaaactta ttattgggtt ggacaggctg gattacatca aagggtgacc tcaaaaactc 1680
378 cagcggttgg aaattttctt ggagagacac cctgagtggg ttggaaaagt tgttttgata 1740
380 cagggtggctg tcccctcacg aggggacgtt gaagaatata aatctttgag ggcagctgta 1800
382 aatgagctag tgggaagaat caatggtaga tttggtaccg tcgaatttgt tccatccat 1860
384 ttccctcata aaagcgtgaa cttccaagag ctgatatctg tctacgctgc tagtgatgtt 1920
386 tgtgtagtgt catcgacacg ggacggaatg aatttggtca gttatgaata cattgcttgt 1980
388 caacaagatc gaaagggatc tctagtacta agtgaatttg cgggagctgc tcagtcatta 2040
390 aatggcgctc tcgtagtgaa tccatggaat acagaagaac tcagtgaagc tatttacgaa 2100
392 ggcttgatca tgagtgaaga gaaaaggagg ggcaattttc agaagatgtt caagtacatt 2160
394 gagaaatata ctgcaagtta ttggggagag aactttgtga aagaattgac gagagtgtga 2220
396 ttactgtggt ttgcaggtta atttgaaaatg ttcacttgta cttgaagaat tttatattat 2280
398 atacatgtta tacatcaata ggataaaaat taagtagaca aagttatcat tttgttgggc 2340
400 tgtaaaaatt gaacgataac aatatatttg acaaaaattaa tttgatctaa ttgagctgga 2400
402 gggcgtaata tatttggttt cctgaatcat cttgtagatc acaatatggg gcagcttctt 2460
404 tcgcagccga tcacagagaa acacatcaca cttgtccaac atgatcacat atcgattca 2520
406 atcggggaaa tgcaaggata caggttgacc atggaagacg cgttctgtga tttgaacgaa 2580
408 agaattattc tgacggaaga gggacttgac atcagaaaac aagacgagaa tacagagggt 2640
410 gatctggagt ctcttcaa at taacatttat ggtgtctttg acggacatgg cgggtt 2695

```

413 <210> SEQ ID NO: 9

414 <211> LENGTH: 26

415 <212> TYPE: DNA

416 <213> ORGANISM: Artificial Sequence

418 <220> FEATURE:

419 <223> OTHER INFORMATION: PCR primer F1 (forward)

422 <220> FEATURE:

423 <221> NAME/KEY: misc_feature

424 <222> LOCATION: (1)..(26)

425 <223> OTHER INFORMATION: n is a, c, t, or g; v is a, c, or g; y is c or t;

427 <400> SEQUENCE: 9

W--> 428 tggccvytnt tccaytacca tccygg

26

431 <210> SEQ ID NO: 10

432 <211> LENGTH: 24

433 <212> TYPE: DNA

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 05/05/2003
PATENT APPLICATION: US/09/927,811C TIME: 14:22:48

Input Set : A:\029474-5006.txt
Output Set: N:\CRF4\05052003\I927811C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 1,5,6,7,8,9,10,11,15

Seq#:3; N Pos. 1,5,6,10,11,15

Seq#:9; N Pos. 9

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/927,811C

DATE: 05/05/2003

TIME: 14:22:48

Input Set : A:\029474-5006.txt

Output Set: N:\CRF4\05052003\I927811C.raw

L:73 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:91 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0